

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:16:37 ; Search time 91.75 Seconds  
(without alignments)  
15.839 Million cell updates/sec

Title: US-09-331-631a-3\_COPY\_29\_73

Perfect score: 252  
Sequence: 1 SEFDROEYECRCQCMQLET.....RCVSOCDKREEDIDMSKYD 45

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77.5	30.8	588	1	VCLB_GOSHI
2	72	28.6	605	1	VCLB_GOSHI
3	67.5	26.8	47	1	AGRP_IDEFCY
4	67	26.6	1170	1	TSP1_HUMAN
5	67	26.6	1170	1	TSP1_MOUSE
6	65	25.8	1173	1	TSP1_XENLA
7	63	25.0	524	1	SRP_SOYBN
8	57	22.6	724	1	AT11_VACCV
9	57	22.6	726	1	AT11_VACCV
10	57	22.6	1284	1	AT11_CAMPX
11	55	21.8	305	1	PPP6_RAT
12	55	21.8	795	1	ENPL_CHICK
13	54.5	21.6	244	1	ENPL_CAEEL
14	54.5	21.6	1021	1	YLB8_CAEEL
15	54	21.4	291	1	CGD1_BRARE
16	54	21.4	643	1	RO60_CAEEL
17	53.5	21.2	301	1	G197_HUMAN
18	53.5	21.2	1170	1	TSP2_BOVIN
19	53	21.0	702	1	AT11_VARV
20	52.5	20.8	91	1	UCR8_HUMAN
21	52.5	20.8	1172	1	TSP2_HUMAN
22	52	20.6	33	1	MBP1_MAIZE
23	52	20.6	37	1	CG23_LUPAN
24	52	20.6	82	1	C2_OXYNO
25	52	20.6	716	1	ENPL_RABIT
26	52	20.6	802	1	ENPL_MOUSE
27	52	20.6	803	1	ENPL_HUMAN
28	52	20.6	804	1	ENPL_CANFA
29	52	20.6	1188	1	T1R_ECOLI
30	51.5	20.4	106	1	COLA_HORSE
31	51.5	20.4	108	1	COLB_HORSE
32	51.5	20.4	147	1	LAMP_HABOF
33	51.5	20.4	298	1	HIS1_CANAL

34	51.5	20.4	525	1	V107_MBTUA	Q57571 methanococ
35	51.5	20.4	563	1	MUC5_BOVIN	P98091 bos taurus
36	51.5	20.4	860	1	U152_HSV6U	P52467 herpes slimp
37	51.5	20.4	860	1	U152_HSV6U	P52540 herpes slimp
38	51.5	20.4	1557	1	DVA1_DICVT	Q24702 dictyocauli
39	51	20.2	418	1	CD15_HUMAN	Q14004 homo sapien
40	51	20.2	741	1	F1BA_CHICK	P14448 gallus galli
41	51	20.2	881	1	ARP8_YEAST	Q12386 saccharomyc
42	51	20.2	893	1	YM92_CAEEL	P34531 caenorhabdi
43	51	20.2	1231	1	CFAN_HUMAN	P08603 homo sapien
44	51	20.2	1370	1	Z261_HUMAN	Q14202 homo sapien
45	50.5	20.0	264	1	RP0D_METTH	O26144 methanobact

## ALIGNMENTS

RESULT	ID	VCLB_GOSHI	STANDARD	PRT	588 AA.
1	AC	P09801			
DT	01-MAR-1989	(Rel. 10, Created)			
DT	01-MAR-1989	(Rel. 10, Last sequence update)			
DT	15-JUL-1999	(Rel. 38, Last annotation update)			
DE	VICILIN C72 PRECURSOR (ALPHA-GLOBULIN B).				
OS	Gossypium hirsutum (Upland cotton).				
OC	Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids II: Malvales; Malvaceae; Gossypium.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Chian C.A., Pyle J.B., Legocki A.B., Dure L. III;				
RT	"Developmental biochemistry of cottonseed embryogenesis and germination. XVIII. cDNA and amino acid sequences of the members of the storage protein families."				
RL	Plant Mol. Biol. 7:475-489(1986).				
CC	-I- FUNCTION: SEED STORAGE PROTEIN.				
CC	-I- SUBCELLULAR LOCATION: COTYLEDONARY MEMBRANE-BOUND VACUOLAR PROTEIN BODIES.				
CC	-I- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN, CONVICILIN, CONGLICININ, ETC.).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL; M16891; AAA33071.1; -				
DR	PIR; A30838; FMCNAB.				
DR	HSSP; P50477; ICAX.				
DR	INTERPRO; IPR001113; -				
DR	PFAM; PF00546; Seedstore_7s; 1.				
KW	Seed storage protein; Signal.				
FT	SIGNAL				
FT	CHAIN				
FT	SEQUENCE				
FT	588 AA: 69729 MW: 63E699B29A8BADBB CRC64:				

Query Match 30.8%; Score 77.5; DB 1; Length 588;

Best Local Similarity 40.5%; Pred. No. 0.059; Mismatches 12; Indels 1; Gaps 1;

Matches 15; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

QY 2 SEFDROEYECRCQCMQLETSGMRRCVSOCDKREED 38

DB 81 EDPORRYECCQECRCQOE-ERORPCQOCRCIKRPFDE 116

RESULT 2  
VCLB\_GOSHI STANDARD: PRT: 605 AA.

```

AC P09799;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VICILIN GC72-A PRECURSOR (ALPHA-GLOBULIN A).
OS Gossypium hirsutum (Upland cotton).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Malvales; Malvaceae; Gossypium.
RN [1]
RP SEQUENCE FROM N.A.
RA Chlan C.A., Borrito K., Kamalay J.A., Dure L., III:
RT "Developmental biochemistry of cottonseed embryogenesis and
RT germination. XIX. Sequences and genomic organization of the alpha
RT globulin (vicilin) genes of cottonseed."
RL Plant Mol. Biol. 9:533-546(1987).
CC -1- FUNCTION: SEED STORAGE PROTEIN.
CC -1- SUBCELLULAR LOCATION: CYTOSOL; MEMBRANE-BOUND VACUOLAR PROTEIN
CC BODIES.
CC -1- SIMILARITY: TO OTHER 7S SEED STORAGE PROTEINS (PHASEOLIN, VICILIN,
CC CONVICILIN, CONGLYCININ, ETC.).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC -----
DR EMBL: M19378; AAA33069.1; -
DR PIR: S06398; S06398.
DR HSSP: P50477; ICAX.
DR INTERPRO: IPR001113; -
DR PFM: PF00546; Seedstore_7s: 1.
KW Seed storage protein; Signal.
FT SIGNAL.
FT CHAIN
FT SEQUENCE 605 AA: 71049 MW: C9DB9371C976953B CRC64;
SO SEQUENCE

Query Match 28.6%; Score 72; DB 1; Length 605;
Best Local Similarity 42.5%; Pred. No. 0.26;
Matches 17; Conservative 9; Mismatches 10; Indels 4; Gaps 3;

QY 1 SEED-ROEYECRQCMQLETSQGMRCVSCQCKRFEEDI 39
DB 33 SEDDPQOREDCRRC-QLETRGQTEQ--DKCEDKSEITL 69

RESULT 3
AGRP_LUFCY
ID AGRP_LUFCY STANDARD: PRT; 47 AA.
AC P56568;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 6.5 KDA ARGININE/GLUTAMATE-RICH POLYPEPTIDE (6.5K-AGRP).
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Cucurbitales; Cucurbitaceae; Luffa.
RN [1]
RP SEQUENCE.
RP TISSUE=SEED;
RX MEDLINE=97357433; PubMed=9214759;
RA Kimura M., Park S.S., Sakai R., Yamasaki N., Funatsu G.;
RT "Primary structure of 6.5k-arginine/glutamate-rich polypeptide from
RT the seeds of sponge gourd (Luffa cylindrica).".
RL Biosci. Biotechnol. Biochem. 61:984-988(1997).
CC -1- FUNCTION: STORAGE PROTEIN WHICH PROVIDES NITROGEN AND CARBON
CC RESERVES DURING GERMINATION AND SEEDLING GROWTH.
CC -1- MASS SPECTROMETRY: MW=5693.39; METHOD=MALDI.

```

```

CC -1- SIMILARITY: SOME, TO 7S SEED STORAGE PROTEINS.
KW Seed storage protein.
FT DISULFID 12 33
FT DISULFID 16 29
SQ SEQUENCE 47 AA: 5698 MW: 588B0EC82273AC05 CRC64;

Query Match 26.8%; Score 67.5; DB 1; Length 47;
Best Local Similarity 41.2%; Pred. No. 0.081;
Matches 14; Conservative 4; Mismatches 15; Indels 1; Gaps 1;

QY 5 ROEYECRQCMQLETS-QGMRCVSCQCKRFEEDI 37
DB 6 RVEYACRVRCQVAEHGVKRRRCQVCEKRLRE 39

RESULT 4
TSP1_HUMAN
ID TSP1_HUMAN STANDARD: PRT; 1170 AA.
AC P07996;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE THROMBOSPONDIN 1 PRECURSOR.
GN THBS1 OR TSP1 OR TSP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
KC TISSUE-ENDOTHELIAL CELLS;
RX MEDLINE=87057617; PubMed=2430973;
RA Lawler J., Hynes R.O.;
RT "The structure of human thrombospondin, an adhesive glycoprotein with
RT multiple calcium-binding sites and homologues with several different
RT proteins."
RL J. Cell Biol. 103:1635-1648(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=89139590; PubMed=2918029;
RA Hennessy S.W., Frazier B.A., Kim D.D., Deckwerth T.L.,
RA Baumgartel D.M., Rotwein P., Frazier W.A.;
RT "Complete thrombospondin mRNA sequence includes potential regulatory
RT sites in the 3' untranslated region."
RL J. Cell Biol. 108:729-736(1989).
RN [3]
RP SEQUENCE OF 1-397 FROM N.A.
RX MEDLINE=87157592; PubMed=3030396;
RA Kobayashi S., Eden-McCutchan F., Framson P., Bornstein P.;
RT "Partial amino acid sequence of human thrombospondin as determined by
RT analysis of cDNA clones: homology to malarial circumsporozoite
RT proteins."
RL Biochemistry 25:8418-8425(1986).
RN [4]
RP SEQUENCE OF 1-374 FROM N.A.
RX MEDLINE=86287276; PubMed=3461443;
RA Dixit V.M., Hennessy S.W., Grant G.A., Rotwein P., Frazier W.A.;
RT "Characterization of a cDNA encoding the heparin and collagen binding
RT domain of human thrombospondin."
RL Proc. Natl. Acad. Sci. U.S.A. 83:5449-5453(1986).
RN [5]
RP SEQUENCE OF 1-166 FROM N.A.
RX MEDLINE=89291870; PubMed=2544587;
RA Laberty C.D., Glerman T.M., Dixit V.M.;
RT "Characterization of the promoter region of the human thrombospondin
RT gene. DNA sequences within the first intron increase transcription."
RL J. Biol. Chem. 264:11222-11227(1989).
RN [6]
RP SEQUENCE OF 1028-1170 FROM N.A.
RA la Fleur M., Jobin C., Gauthier J., Kreis C.G.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND
CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,

```

CC LAMININ AND TYPE V COLLAGEN.  
 CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 WMFC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS (WHICH BIND CALCIUM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M25631; AAA36741.1; -  
 CC EMBL: X04665; CAA28370.1; -  
 CC EMBL: X14787; CAA32889.1; -  
 CC EMBL: J04835; AAA61178.1; -  
 CC EMBL: M99425; AAB59366.1; -  
 CC PIR: A05172; A05172.  
 CC PIR: A25812; A25812.  
 CC PIR: A26155; A26155.  
 CC PIR: A30140; A30140.  
 CC PIR: A34274; A34274.  
 CC HSSP: P35555; 1EMO.  
 CC MIM: 188060; -  
 CC INTERPRO: IPR000561; -  
 CC INTERPRO: IPR000884; -  
 CC INTERPRO: IPR001007; -  
 CC PFAM: PF00008; EGF; 2.  
 CC PFAM: PF00090; tsp.1; 3.  
 CC PFAM: PF00093; vwc.1; 1.  
 CC PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
 CC PROSITE: PS01186; EGF\_2; 1.  
 CC PROSITE: PS00092; TSP1; 3.  
 CC PROSITE: PS01208; WMFC; 1.  
 CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 CC EGF-like domain; Signal.  
 CC SIGNL 1 18  
 CC CHAIN 19 1170  
 CC DOMAIN 19 232  
 CC DOMAIN 316 373  
 CC DOMAIN 379 548  
 CC DOMAIN 549 690  
 CC DOMAIN 723 950  
 CC DOMAIN 951 1170  
 CC REPEAT 379 430  
 CC REPEAT 435 491  
 CC REPEAT 492 548  
 CC DOMAIN 548 587  
 CC DOMAIN 588 645  
 CC DOMAIN 646 690  
 CC REPEAT 723 759  
 CC REPEAT 759 781  
 CC REPEAT 782 840  
 CC REPEAT 841 878  
 CC REPEAT 879 914  
 CC REPEAT 915 950  
 CC SITE 926 928  
 CC DISULFID 270 270  
 CC DISULFID 274 274  
 CC DISULFID 551 562  
 CC DISULFID 556 572  
 CC DISULFID 575 586  
 CC DISULFID 592 608  
 CC DISULFID 599 617  
 CC DISULFID 620 644  
 CC DISULFID 650 663  
 CC DISULFID 657 676  
 CC DISULFID 678 689

FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 84 84 T -> A (TN REF. 2, 3 AND 4).  
 FT CONFLICT 523 523 T -> A (TN REF. 2).  
 SQ SEQUENCE 1170 AA: 129412 MW: 69B3E5EAE3A395E CRC64;  
 Query Match 26.6%; Score 67; DB 1; Length 1170;  
 Best Local Similarity 39.5%; Pred. No. 1.9;  
 Matches 15; Conservative 6; Mismatches 11; Indels 6; Gaps 2;  
 QY 13 ROCMOLE-----TSGOMRRC-VSQCCKRFEDIDWSEKY 44  
 DB 404 RSCDSLNRCGSSVQTRCHIECKRRKQDGNSHW 441  
 RESULT 5  
 TSP1\_MOUSE  
 ID TSP1\_MOUSE STANDARD; PRT: 1170 AA.  
 AC P35441;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE THROMBOSPONDIN 1 PRECURSOR.  
 GN THBS1 OR TSP1.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92128941; PubMed-1774063;  
 RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,  
 RT Jenkins N.A.;  
 RT "Characterization of the murine thrombospondin gene";  
 RL Genomics 11:587-600(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92147683; PubMed-1371115;  
 RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,  
 RA Dixit V.M.;  
 RT "Characterization of mouse thrombospondin 2 sequence and expression  
 RT during cell growth and development";  
 RL J. Biol. Chem. 267:3274-3281(1992).  
 RN [3]  
 RP SEQUENCE OF 1-490 FROM N.A.  
 RX MEDLINE-90375546; PubMed-2398070;  
 RA Bornstein P., Alfi D., Devarayalu S., Framson P., Li P.;  
 RT "Characterization of the mouse thrombospondin gene and evaluation of  
 RT the role of the first intron in human gene expression";  
 RL J. Biol. Chem. 265:16691-16698(1990).  
 CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND  
 CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,  
 CC LAMININ AND TYPE V COLLAGEN.  
 CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 WMFC DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS  
 CC (WHICH BIND CALCIUM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M62470; AAA50611.1; -  
 CC EMBL: M62450; AAA50611.1; JOINED.  
 DR EMBL: M62451; AAA50611.1; JOINED.

DR EMBL: M62452; AAAS0611.1; JOINED.  
 DR EMBL: M62453; AAAS0611.1; JOINED.  
 DR EMBL: M62454; AAAS0611.1; JOINED.  
 DR EMBL: M62455; AAAS0611.1; JOINED.  
 DR EMBL: M62456; AAAS0611.1; JOINED.  
 DR EMBL: M62457; AAAS0611.1; JOINED.  
 DR EMBL: M62458; AAAS0611.1; JOINED.  
 DR EMBL: M62459; AAAS0611.1; JOINED.  
 DR EMBL: M62460; AAAS0611.1; JOINED.  
 DR EMBL: M62461; AAAS0611.1; JOINED.  
 DR EMBL: M62462; AAAS0611.1; JOINED.  
 DR EMBL: M62463; AAAS0611.1; JOINED.  
 DR EMBL: M62464; AAAS0611.1; JOINED.  
 DR EMBL: M62465; AAAS0611.1; JOINED.  
 DR EMBL: M62466; AAAS0611.1; JOINED.  
 DR EMBL: M62467; AAAS0611.1; JOINED.  
 DR EMBL: M62468; AAAS0611.1; JOINED.  
 DR EMBL: M62469; AAAS0611.1; JOINED.  
 DR EMBL: M87276; AAAS3063.1; -.  
 DR EMBL: J05606; AAA40431.1; -.  
 DR EMBL: J05605; AAA40431.1; JOINED.  
 DR PIR: B42587; B42587.  
 DR PIR: A37905; A37905.  
 DR HSSP: P35555; 1EMO.  
 DR MGD: MGI:98737; THBS1.  
 DR INTERPRO: IPR000561; -.  
 DR INTERPRO: IPR000884; -.  
 DR INTERPRO: IPR001007; -.  
 DR PFAM: PF000008; EGF\_2.  
 DR PFAM: PF000090; tsp\_1; 3.  
 DR PFAM: PF00093; vwc; 1.  
 DR PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS50092; TSP1; 3.  
 DR PROSITE: PS01208; VMFC; 1.  
 DR Glycoprotein: Cell adhesion; Calcium-binding; Heparin-binding; Repeat; EGF-like domain; Signal.  
 KM SIGNAL 1 18  
 FT CHAIN 19 1170 THROMBOSPONDIN 1.  
 FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL);  
 FT DOMAIN 316 373 VMFC.  
 FT DOMAIN 379 548 3 X TSP TYPE-1 REPEATS (CS-LIKE).  
 FT DOMAIN 549 690 3 X EGF-TYPE REPEATS.  
 FT DOMAIN 723 950 7 X TSP TYPE-3 REPEATS (CA-BINDING).  
 FT DOMAIN 951 1170 C-TERMINAL.  
 FT REPEAT 379 430 TSP TYPE-1 1.  
 FT REPEAT 435 491 TSP TYPE-1 2.  
 FT REPEAT 492 548 TSP TYPE-1 3.  
 FT DOMAIN 549 587 EGF-LIKE 1.  
 FT DOMAIN 588 645 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT DOMAIN 646 690 EGF-LIKE 3.  
 FT REPEAT 723 758 TSP TYPE-3 1.  
 FT REPEAT 759 781 TSP TYPE-3 2.  
 FT REPEAT 782 817 TSP TYPE-3 3.  
 FT REPEAT 818 840 TSP TYPE-3 4.  
 FT REPEAT 841 878 TSP TYPE-3 5.  
 FT REPEAT 879 914 TSP TYPE-3 6.  
 FT REPEAT 915 950 TSP TYPE-3 7.  
 FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).  
 FT DISULFID 270 270 INTERCHAIN (PROBABLE).  
 FT DISULFID 274 274 INTERCHAIN (PROBABLE).  
 FT DISULFID 551 562 BY SIMILARITY.  
 FT DISULFID 556 572 BY SIMILARITY.  
 FT DISULFID 575 586 BY SIMILARITY.  
 FT DISULFID 592 608 BY SIMILARITY.  
 FT DISULFID 609 617 BY SIMILARITY.  
 FT DISULFID 620 644 BY SIMILARITY.  
 FT DISULFID 650 663 BY SIMILARITY.  
 FT DISULFID 657 676 BY SIMILARITY.  
 FT CARBOHYD 678 689 BY SIMILARITY.  
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 708 708 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1067 1067 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 1025 1025 F -> L (IN AAAS3063).  
 SQ SEQUENCE 1170 AA; 129646 MW; 0443E493615E7F06 CRC64;  
 Query Match 26.6%; Score 67; DB 1; Length 1170;  
 Best Local Similarity 39.5%; Pred. No. 1.9;  
 Matches 15; Conservative 6; Mismatches 11; Indels 6; Gaps 2;  
 QY 13 ROCMOLE-----TSGOMRRC-VSQCDKRFREEDIDMSKY 44  
 Db 404 RSCDSLNNRCESSVQTRCHIQECDKRFKDGGSW 441  
 RESULT 6  
 TSPL\_XENLA  
 ID TSPL\_XENLA STANDARD; PRT: 1173 AA.  
 AC P33448;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE THROMBOSPONDIN 1 PRECURSOR.  
 GN THBS1 OR TSPL.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC xenopodinae; Xenopus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;  
 RA Submitted (xx-1993) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: ADHESIVE GLYCOPROTEIN THAT MEDIATES CELL-TO-CELL AND  
 CC CELL-TO-MATRIX INTERACTIONS. CAN BIND TO FIBRINOGEN, FIBRONECTIN,  
 CC LAMININ AND TYPE V COLLAGEN.  
 CC -1- SUBUNIT: HOMOTRIMER, CROSS-LINKED BY DISULFIDE BONDS.  
 CC -1- SIMILARITY: BELONGS TO THE THROMBOSPONDIN FAMILY.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 TYPE-1 TSP REPEATS AND 7 TYPE-3 TSP REPEATS  
 CC (WHICH BIND CALCIUM).  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L04278; -; NOT\_ANNOTATED\_CDS.  
 DR HSSP: P00740; 11XA.  
 DR INTERPRO: IPR000561; -.  
 DR INTERPRO: IPR000884; -.  
 DR INTERPRO: IPR001007; -.  
 DR PFAM: PF00008; EGF\_2.  
 DR PFAM: PF00090; tsp\_1; 3.  
 DR PFAM: PF00093; vwc; 1.  
 DR PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS50092; TSP1; 3.  
 DR PROSITE: PS01208; VMFC; 1.  
 DR Glycoprotein: Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 KM EGF-like domain; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 1173 THROMBOSPONDIN 1.  
 FT DOMAIN 23 235 HEPARIN-BINDING (POTENTIAL).  
 FT DOMAIN 319 376 VMFC.  
 FT DOMAIN 382 546 3 X TSP TYPE-1 REPEATS (CS-LIKE).  
 FT DOMAIN 550 693 3 X EGF-TYPE REPEATS.  
 FT DOMAIN 724 953 7 X TSP TYPE-3 REPEATS (CA-BINDING).  
 FT DOMAIN 954 1173 C-TERMINAL.  
 FT REPEAT 382 433 TSP TYPE-1 1.

```

FT REPEAT 438 494 TSP TYPE-1 2.
FT REPEAT 495 546 TSP TYPE-1 3.
FT DOMAIN 550 590 EGF-LIKE 1.
FT DOMAIN 591 648 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 649 693 EGF-LIKE 3.
FT REPEAT 726 761 TSP TYPE-3 1.
FT REPEAT 762 784 TSP TYPE-3 2.
FT REPEAT 785 820 TSP TYPE-3 3.
FT REPEAT 821 843 TSP TYPE-3 4.
FT REPEAT 844 881 TSP TYPE-3 5.
FT REPEAT 882 917 TSP TYPE-3 6.
FT REPEAT 918 953 TSP TYPE-3 7.
FT SITE 929 931 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 554 565 BY SIMILARITY.
FT DISULFID 559 575 BY SIMILARITY.
FT DISULFID 578 589 BY SIMILARITY.
FT DISULFID 595 611 BY SIMILARITY.
FT DISULFID 602 620 BY SIMILARITY.
FT DISULFID 623 647 BY SIMILARITY.
FT DISULFID 660 679 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 705 705 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 711 711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1070 1070 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1173 AA: 130019 MW: 199036D6516C0F24 CRC64;

```

```

Query Match 25.8%; Score 65; DB 1; Length 1173;
Best Local Similarity 39.5%; Pred. No. 3 2;
Matches 15; Conservative 5; Mismatches 12; Indels 6; Gaps 2;

```

```

OY 13 RQCMQLE-----TSGQMRRC-VSQCDKFEEDIDWSKY 44
      | | | | | | | | | | | | | | | | | |
Db 407 RSCDSLNNPCGSSVQTRSCQIDCDKRFKQDGMSHW 444

```

```

RESULT 7
SBP_SQYBN STANDARD; PRT; 524 AA.
AC 004672;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE SUCROSE-BINDING PROTEIN PRECURSOR (SBP).
GN SBP.
OS Glycine max (Soybean).
OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
OC Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eurosids I:
OC Fabales: Fabaceae; Papilionoideae; Glycine.
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 30-50.
RC TISSUE-EMBRYO;
RX MEDLINE-93104680; PubMed-1467654;
RA Grimes H.D., Overvoorde P.J., Ripp K., Franceschi V.R., Hitz W.D.;
RT "A 62-kD sucrose binding protein is expressed and localized in
RT tissues actively engaged in sucrose transport.";
RL Plant Cell 4:1561-1574(1992).
CC - FUNCTION: PLAYS A ROLE IN SUCROSE TRANSPORT.
CC - SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED.
CC - TISSUE SPECIFICITY: ASSOCIATED WITH THE PLASMA MEMBRANE OF
CC SEVERAL CELL TYPES ENGAGED IN SUCROSE TRANSPORT, INCLUDING THE
CC MESOPHYLL CELLS OF YOUNG SINK LEAVES, THE COMPANION CELLS OF
CC MATURE PHLOEM AND THE CELLS OF DEVELOPING COTYLEDONS.
CC - DEVELOPMENTAL STAGE: IN THE COTYLEDON, EXPRESSION IS NOT DETECTED
CC UNTIL 10 DAYS AFTER FERTILIZATION, BETWEEN 10-19 DAYS AFTER
CC FERTILIZATION, EXPRESSION INCREASES RAPIDLY BUT DECLINES 20-30
CC DAYS AFTER FERTILIZATION, 30 DAYS AFTER FERTILIZATION, NO
CC EXPRESSION OCCURS. THIS EXPRESSION PATTERN CLOSELY PARALLELS THE

```

```

CC RATE OF SUCROSE UPTAKE IN THE COTYLEDON.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L06038; AAB03894.1; -.
DR PIR; J01730; J01730.
DR HSSP; P50477; ICAM.
DR INTERPRO; IPR001113; -.
DR PPM; PF00546; Seedstore-7s; 1.
KW transport; Sugar transport; Signal; Membrane.
FT SIGNAL 1 29
FT CHAIN 30 524
SQ SEQUENCE 524 AA: 60522 MW: 0251EE90796EF341 CRC64;

```

```

Query Match 25.0%; Score 63; DB 1; Length 524;
Best Local Similarity 38.2%; Pred. No. 2 6;
Matches 13; Conservative 4; Mismatches 15; Indels 2; Gaps 1;

```

```

OY 2 EFDROEYEECKRQCMQLE--TSGQMRRCVSCDK 33
      | | | | | | | | | | | | | | | |
Db 34 EEDPELVYCKHCCQCCQOQYTTEDKRVCLQSCDR 67

```

```

RESULT 8
ATIL_VACCV STANDARD; PRT; 724 AA.
AC P24759;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE 94 KDA A-TYPE INCLUSION PROTEIN (ATI).
OS Vaccinia virus (strain WR).
OC viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-91310644; PubMed-1856205;
RA Amegadzie B.Y., Ahn B.-Y., Moss B.;
RT "Identification, sequence, and expression of the gene encoding a Mr
RT 35,000 subunit of the vaccinia virus DNA-dependent RNA polymerase.";
RL J. Biol. Chem. 266:13712-13718(1991).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE-92124754; PubMed-1733111;
RA Amegadzie B.Y., Sisler J.R., Moss B.;
RT "Frame-shift mutations within the vaccinia virus A-type inclusion
RT protein gene.";
RL Virology 186:777-782(1992).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE-92074241; PubMed-1962448;
RA de Carlos A., Paes E.;
RT "Isolation and characterization of mutants of vaccinia virus with a
RT modified 94-kDa inclusion protein.";
RL Virology 185:768-778(1991).
CC - FUNCTION: MATURE VIRUSES ARE OCCURRED INTO THE ATI, AND IT HAS
CC BEEN ASSUMED THAT SUCH BODIES PROTECT THE VIRUS DURING
CC DISSEMINATION FROM ANIMAL TO ANIMAL.
CC - I- MISCELLANEOUS: A CHARACTERISTIC FEATURE OF ATI IS THE FORMATION OF
CC LARGE MASSES WITH NO SURROUNDING MEMBRANES IN THE CYTOPLASM OF
CC INFECTED CELLS.
CC - SIMILARITY: 92% IDENTITY TO COMPOX VIRUS A-TYPE INCLUSION
CC PROTEIN N-TERMINAL.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```





```

FT CONFLICT 206 206 D -> E (IN REF. 3 AND 4).
FT CONFLICT 267 267 V -> L (IN REF. 3).
FT CONFLICT 303 303 E -> Q (IN REF. 3).
FT CONFLICT 307 307 N -> D (IN REF. 3 AND 4).
FT CONFLICT 317 317 E -> H (IN REF. 3).
FT CONFLICT 378 378 C -> A (IN REF. 3).
FT CONFLICT 593 593 E3 -> DR (IN REF. 3).
FT CONFLICT 653 653 W -> C (IN REF. 3).
FT CONFLICT 669 675 GMDISTN -> VFSS (IN REF. 4).
SQ SEQUENCE 795 AA: 91555 MW: BEIB29E1DBEC5A9A CRC64:

```

```

Query Match 21.8%; Score 55; DB 1; Length 795;
Best Local Similarity 23.8%; Pred. No. 32;
Matches 10; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

```

```

OY 2 EFDROEYECRQCMQLETSQGMRCVSCDCKRFEEDIDMSK 43
DB 581 EFDGKRFQNVAREGVAFEESEKSKESREALEKEFEPLNMMK 622

```

## RESULT 13

```

YL53_CAEEL STANDARD: PRT: 244 AA.
AC SW4433:
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE HYPOTHETICAL 29.0 KDA PROTEIN F44E2.3 IN CHROMOSOME III.
GN F44E2.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodetinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Lalister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smailton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Spoat J.,
RA Woldman P.;
RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.*;
RL Nature 368:32-38(1994).
CC -i- SIMILARITY: TO DNAJ.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

CC EMBL: U23646; AAA28041.1; -.
DR PIR: S44822; S44822.
DR MORMPER: F44E2.3; CE00181.
KW Hypothetical protein.
FT DOMAIN 3 45 ARG/ASP/LYS-RICH.
FT DOMAIN 79 90 PRO-RICH.
SO SEQUENCE 244 AA: 28994 MW: C0CE677FE01A2B18 CRC64:

```

```

Query Match 21.6%; Score 54.5; DB 1; Length 244;
Best Local Similarity 39.4%; Pred. No. 12;
Matches 13; Conservative 8; Mismatches 7; Indels 5; Gaps 2;

```

```

OY 9 EBCRRCQCMQLETSQGMRCVSCD---KRFED 38
DB 164 EBSRKKCKQLE--AELEKRVLEAESRKKFEED 194

```

## RESULT 14

```

YL88_CAEEL STANDARD: PRT: 1021 AA.
ID YL88_CAEEL
AC P46582:
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHETICAL 111.9 KDA PROTEIN C34E10.8 IN CHROMOSOME III.
GN C34E10.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodetinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Kirsten J.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----

```

```

DR EMBL: U10402; AAA19070.1; -.
DR MORMPER: C34E10.8; CE01188.
KW Hypothetical protein.
SQ SEQUENCE 1021 AA: 111858 MW: F1B63DAA26332F5F CRC64:

```

```

Query Match 21.6%; Score 54.5; DB 1; Length 1021;
Best Local Similarity 22.0%; Pred. No. 46;
Matches 11; Conservative 15; Mismatches 13; Indels 11; Gaps 2;

```

```

OY 1 SEFDROEYECR-----QCMQLETSQGMRCVSCDCKRFEEDIDMSK 43
DB 24 NEYDKRYGQCMYASYRTLVRCKQIRSKELAKHGRC----EHVFSK 69

```

## RESULT 15

```

CGD1_BRARE STANDARD: PRT: 291 AA.
AC Q90459:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE G1/S-SPECIFIC CYCLIN D1.
GN CYCD1.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Rasbortinae; Danio.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96138542; PubMed=8547308;
RA Yarden A., Salomon D., Geiger B.;
RT "Zebrafish cyclin D1 is differentially expressed during early
RT embryogenesis.";
RL Biochim. Biophys. Acta 1264:257-260(1995).
CC -i- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G1/S
CC (START) TRANSITION.
CC -i- SUBUNIT: INTERACTS WITH THE CDK4 AND CDK6 PROTEIN KINASES TO FORM
CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT
CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.
CC -i- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN D SUBFAMILY.

```



```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC
DR EMBL: X87581; CAA60885.1; "-
DR ZFIN: ZDB-GENE-980526-176; CYCD1.
DR INTERPRO: IPR000553; "-
DR PfAM: PF00134; cyclin; 1.
DR PROSITE: PS00292; CYCLINS; 1.
DR Kyr: Cyclin; Cell cycle; Cell division.
DR SEQUENCE 291 AA; 33067 MW; FA5274C8B1B46D5EF CRC64;

```

Query Match	21.48;	Score 54;	DB 1;	Length 291;
Best Local Similarity	38.58;	Pred. No. 16;		
Matches 15;	Conservative 7;	Mismatches 11;	Indels 6;	Gaps 2;

```

QY      10 ECRQCMQ-----LETS-GCMRCVSCDKRFEEDIDMS 42
      : | | : || | : : : | | | | |
Db      242 DCLRSCQEQIESLLESSLRQAQGHISTETKRVEEVDVLS 280

```

```
Search completed: March 1, 2001, 16:16:38
Job time: 418 sec
```

